Education data: what's the problem?

How can schools make good decisions about EdTech and the data it collects? Sonia Livingstone and Kruakae Pothong interviewed Ben Williamson about the growing influence of Big Tech in public education systems, and how this has been accelerated by the move to remote learning during COVID-19.

Sonia: We're interested in the role of commercial interests in shaping EdTech developments.

Ben: In our report for Education International we recognised that commercial organisations were hugely important during COVID-19 in ensuring continuity of learning for 1.6 billion students around the world whose schools or colleges were closed. It's about trying to strike a balance, to ensure that there aren't unforeseen side-effects of enabling the private sector to provide more and more of the infrastructural supports to schooling.

Kruakae: What do you see as the problem?

Ben: We need to understand what happens when you introduce education technologies, particularly the ones that extract a lot of data from students. Does it change the way that teachers view students or the way students view themselves? When as a child you are encouraged by your teachers and by your school to adopt a kind of data gaze.

Photo by August de Richelieu from Pexels

Sonia: Can schools make meaningful decisions about EdTech or is the power over their heads?

Ben: I suspect there are some schools who have quite strongly worked out internal policies for how they acquire and use different educational technologies and some for whom that's much less developed. I think the challenge, for example with free apps for teachers, is that the model now is direct to consumer where the consumer is either the teacher or indeed the student or parent.

Sonia: Can you give us an example of what EdTech decisions bypass the school entirely, and go straight to the parent?

Ben: You could think of Class Dojo as prototypical example. It's available for free, like any social media app. The teacher signs up to it. The marketing was done through Twitter and Facebook and Instagram. Teachers, certainly in the early stages were starting to use Class Dojo without necessarily going through lots of reports of parental consent not being sought.

To what extent do these things now flow straight to the teacher rather than going through some kind of procurement and vetting process? In some local authorities, they've had to retrospectively come up with procedures and rules about vetting things, and they now have lists of apps that they've said that teachers can start using and others that they don't recommend.

Sonia: So while we may think that companies push their commercial interests, there is also a question about the responsibility of the public sector to develop and update its procurement policies so they are fit for purpose.

Ben: Yes, this debate is playing out in higher education around exam proctoring. In the USA, the argument is that universities and public institutions have not exercised their ethical responsibilities or their oversight and have rushed to buy into the commercial service which came with all the promises of making sure that students now forced to take their exams online weren't cheating. And it has all sorts of effects. Reports of heightened student anxiety. Lots of false positives. Members of staff being sued.

Sonia: But do schools have any genuine say in determining the data processing that they've become part of?

Ben: Google and Microsoft now have structural dominance in schools. That was at least partially enabled by Department for Education funding in England over recent months with a partnership to support schools to roll out with G Suite or Microsoft 365 as a way of ensuring connectivity.

Something like G Suite provides everything. It is in a way a total solution learning management system where it handles your email, it handles your messaging, you can do your PowerPoints through Classroom, you can set work. As a teacher, you can receive work. You can view student analytics. But the actual use is really uneven. A lot of teachers don't access it. So, we shouldn't fall for the hype.

There's also the question of whether a lot of these supposedly cutting-edge analytic systems even do what they're supposed to. The evidence of effectiveness claimed by commercial providers is very rarely openly or transparently available for us to actually look at.

Kruakae: How can schools decide what Edtech services they should be using or what type of data they should be gathering and to what purposes?

Ben: It could be said that schools and teachers are left to the Wild West of the market with thousands of suppliers all marketing to them. Some of those products might be really good, but at the same time, these aren't entirely innocent or impartial or independent organisations.
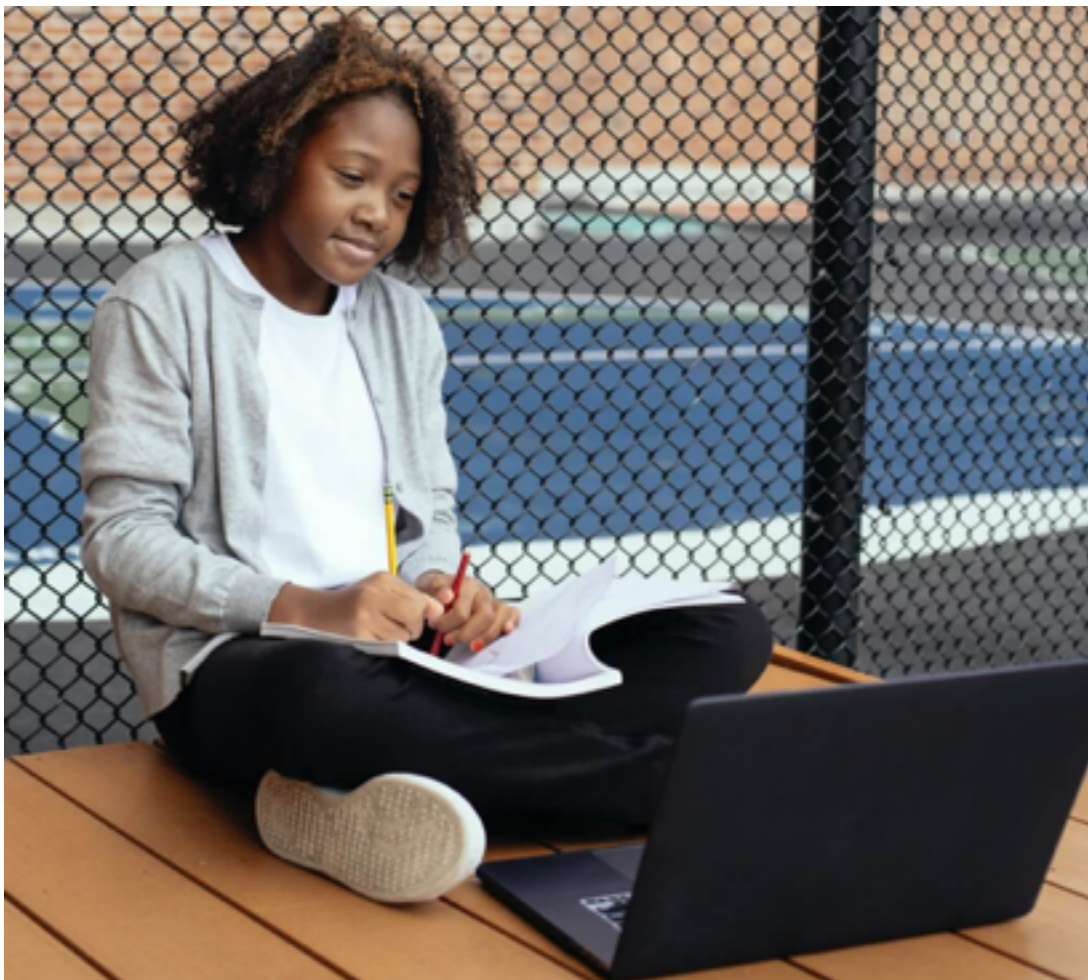
Kruakae: What kind of data is currently being collected in the education space?

Ben: There's a wide variety. You've got the National Pupil Database in England, which is everything official on a child. Their entire school records throughout their lives which can then be linked with higher education data. So, you have a kind of really detailed data fingerprint from the early years through to graduation.

Then you've got the learning management system. If it has learning analytics running in it and is being used in the classroom, then more or less every tap of the keyboard or stroke of the mouse will generate some kind of data point. When we compare them to thousands of others who were on a similar course following the same curriculum, we can see that this student is similar to this cluster here and make a prediction about this individual in terms of that cluster, and that leads to some kind of intervention, allegedly.

The potential for big data generation in education is now happening through the big tech companies like Google. Google Classroom is being used by over 100 million students, and the wider G Suite of apps and services is used by 150 million students. Schools are actively being told to make use of Google Classroom at least in the short term to make sure that if kids all get sent home because of renewed lockdown measures, they can at least access some kind of virtual learning.

Although Google promises that any data collected inside G Suite is never given away for advertisers, it doesn't say what it actually does with it in terms of product development, which is where all the value is. Also, they acknowledge that the second you link, as a child user at a school or home learning, from one of the G Suite apps or platforms into something else say YouTube, then the same privacy policy does not apply.

Google certainly is pitching itself more like a learning management system, the entire underlay for everything digital school wants or needs to do. Rather than Google saying, look we've got a cool video service which you can use in your teaching, it's saying actually, we are the underlay. School closures were the ideal opportunity to roll that out and enable a much larger population of teachers and students and parents to experience it.

Sonia: So Ben, I'm assuming you wouldn't take us back to the 19th century, so would you hold out hopes of better design for EdTech? Or is better legislation needed? Or is this a data literacy problem and everyone should wise up and make better choices?

Ben: I think there's a good argument for some data literacy stuff but for teachers and school leaders and children. But I recognise that is incredibly tricky and also it's a partial solution. Some form of regulation is necessary, perhaps so that privately produced data that is extracted from public services should be made available to the public for alternative forms of use. I'm also leaning towards a properly independent body to have oversight of this sector. What we saw during lockdown was a rush to Google and Microsoft where things could've been otherwise. There were lots of people arguing there are all of these open educational resources that have been developed by organisations that are specifically trying to advance education without it being an entirely commercial enterprise.

Ben Williamson is a Chancellor's Fellow at the Centre for Research in Digital Education and the Edinburgh Futures Institute, examining the intersections of digital technologies, science, and data with education policy and governance. His current research focuses on two key themes: 1) the expansion of educational data infrastructures to enable information to be collected from schools and universities, then analysed and circulated to various audiences, 2) the emergence of 'intimate data' relating to students' psychological states, neural activity, and genetic profiles, and the implications for increasingly scientific ways of approaching educational policy and practice.